



SEGMENTATION D'IMAGES PAR COMPARAISON D'HISTOGRAMMES DE LOIS D'INNOVATIONS

P. COURTELLEMONT, C. OLIVIER, O. COLOT & A. EL MATOUAT*

La3i-LACIS - UFR des Sciences - Université de Rouen
76821 Mont Saint-Aignan Cédex, FRANCE
Tel: 35 14 68 74, Fax: 35 14 63 49

*Ecole Normale Supérieure
BP 2437, FEZ Principale, MAROC

RÉSUMÉ

ABSTRACT

Nous proposons, dans ce papier, une méthode de détection de changements de modèles à partir de modélisations de type AR par la comparaison d'histogrammes approchant des lois de probabilité dont le nombre de classes est déterminé par différents critères d'information. Ces critères ont été initialement développés pour la sélection d'ordre de modèles ou pour des techniques de codage, et sont ici reformulés pour être appliqués aux histogrammes. Le support est constitué de fichiers images issus de photographies comportant des zones de caractères à localiser. Il s'agit de différencier les zones de fond des zones de caractères par comparaisons des histogrammes de leur loi d'innovation respective.

In this paper, we propose a method for detecting model changes (from AR modelling) by comparison of histograms approximating probability laws. The number of bins of these histograms is computed thanks to different information criteria. These criteria had initially been developed for model order selection or for coding. They have been rewritten to be applied to histograms.

This study deals with photographs including written areas to be located. Background areas have to be distinguished from characters areas by comparing the histograms of their innovation laws respectively.

1. INTRODUCTION

Le support de l'étude est constitué d'erreurs de modélisation d'images de 512x512 pixels, codés sur 256 niveaux de gris. Ces images, issues de photographies, comportent des zones de caractères à localiser. Les images des erreurs, issues d'une modélisation AR, servent de support à différentes études entreprises au sein du La3i-LACIS comme la recherche d'axes de lecture par transformée de Hough [COUR92]...

Le but de ce papier est de proposer une méthode de localisation des changements de lois des erreurs de prédiction.

Dans une première partie, nous décrivons comment ces images d'erreurs ont été obtenues. Une modélisation autorégressive s'appuyant sur l'hypothèse, validée sur les exemples traités, d'un tracé de fond correctement modélisable par un AR 2-D d'ordre pxq peu élevé est appliquée. Pour chaque pixel de l'image, l'erreur de prédiction en niveau de gris ε_k est calculée et comparée à un seuil dont le dépassement enclenchera une procédure de reconnaissance de rupture. La probabilité de dépassement du seuil est connue car sur le tracé de fond, les (ε_k) sont de loi reconnue gaussienne. Lorsque le seuil est franchi, le bloc des N pixels suivants est analysé de façon à identifier la nature de l'anomalie.

Les N erreurs de prédiction calculées sur cette fenêtre constituent le N -échantillon dont la loi va être approchée par un histogramme de classes de pas variable; la construction de cet histogramme fera l'objet de la deuxième partie de ce papier. Cet histogramme doit estimer de manière optimale au sens du maximum de vraisemblance, la loi réelle λ suivie par les ε_k . Il s'agit de déterminer le nombre de classes de l'histogramme à

partir de critères, issus de coûts au sens du maximum de vraisemblance, dont l'expression sera toujours de la forme:

$$f(c, N) = -2 \sum_{i=1}^c \theta_c(B_i) \text{Log} \frac{\theta_c(B_i)}{\mu(B_i)}$$

où les B_i forment une partition de l'histogramme cherché à c classes, $\theta_c(B_i)$ la loi de probabilité des ε_k obtenue par cet histogramme et μ une loi a priori fixée.

Pour chaque critère, le nombre c retenu est celui qui minimise ce critère. Il est obtenu par un algorithme de fusion de classes adjacentes.

Dans une troisième partie, nous abordons la segmentation de l'image en zones homogènes. Elle se fait par comparaison d'histogrammes des lois d'innovation du tracé de fond et de la zone repérée. La comparaison se fait par calcul de mesures de dissemblance entre les histogrammes au nombre de classes déterminé dans la deuxième partie.

Nous présentons dans une troisième partie, les résultats obtenus sur un type d'image avec texte: le comportement des critères, leur comparaison, les mesures de dissemblance lors des ruptures.

2. IMAGES TRAITÉES

2.1. Modélisation AR

Rappelons les notations bien connues suivantes: on considère les modèles AR de signaux:



$$y_k = - \sum_{i=1}^p a_i y_{k-i} + e_k \quad (2.1)$$

où y_k est le signal échantillonné en k , e_k un bruit blanc supposé gaussien et les a_i les paramètres du modèle d'ordre p . Si on note $A^T = (-a_1, \dots, -a_p)$ le vecteur transposé des paramètres, et Y_{k-1}^T le vecteur transposé des observations passées: $Y_{k-1}^T = (y_{k-1}, \dots, y_{k-p})$, l'équation (2.1) s'écrit alors:

$$y_k = A^T Y_{k-1} + e_k \quad (2.2)$$

On cherche à identifier le modèle suivant la méthode des MCR.

Si A_k est la matrice estimée de A à la position k , on obtient alors la double récurrence bien connue suivante:

$$F_{k-1}^{-1} = F_{k-1}^{-1} + Y_{k-1} Y_{k-1}^T \quad (2.3)$$

$$A_k = A_{k-1} + F_{k-1}^{-1} Y_{k-1} \varepsilon_k \quad (2.4)$$

avec : $\varepsilon_k = y_k - z_k$, l'erreur de prédiction a priori, z_k étant le signal estimé en k .

2.2 Extension à une modélisation 2-D

Si on applique la technique précédente aux images, l'observation y_k est le niveau de gris du pixel à la position k sur une ligne d'image. La figure 1 montre un exemple d'image étudiée.

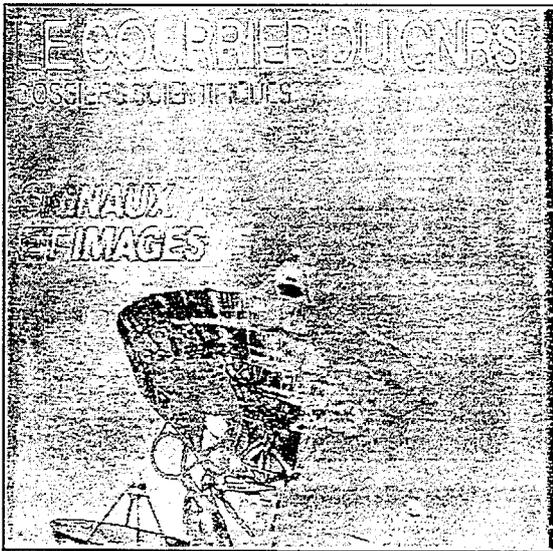


Figure 1: Image traitée 512x512 pixels

Les images test contiennent 512x512 pixels codés sur 256 niveaux de gris définissant les observations $y(m,n)$, $0 \leq m,n \leq 511$. Ces images peuvent provenir de couvertures de magazine (pour l'étude de systèmes de lecture automatique pour non-voyants), de photographies de conteneurs ou de véhicules (dans le but d'identification automatique d'immatriculations), ou d'enveloppes postales de couleur (pour une lecture automatique du libellé).

Soit ainsi une image de dimension $N \times M$ pixels. Un modèle AR 2-D, donnant la valeur $y(m,n)$ du niveau de gris d'un pixel d'abscisse m et d'ordonnée n est défini par:

$$y(m,n) = \sum_{i=0}^{p-1} \sum_{j=0}^{q-1} a(i,j) y(m-i,n-j) + e(m,n) \text{ avec: } i+j \neq 0 \quad (2.5)$$

ou, avec une forme matricielle:

$$y(m,n) = [A_0^T, \dots, A_{p-1}^T] \begin{bmatrix} Y_0 \\ \dots \\ Y_{p-1} \end{bmatrix} + e(m,n) \quad (2.6)$$

avec, si $i \neq 0$:

$$A_i^T = [a(i, 0), \dots, a(i, q-1)], \quad Y_i^T = [y(m-i, n), \dots, y(m-i, n-q+1)]$$

et, si $i=0$:

$$A_0^T = [a(0, 1), \dots, a(0, q-1)], \quad Y_0^T = [y(m, n-1), \dots, y(m, n-q+1)].$$

On écrira (2.6) plus simplement comme:

$$y(m,n) = A^T Y + e(m,n). \quad (2.7)$$

La condition $i+j \neq 0$ permet d'exclure la valeur $y(m,n)$ qui est la valeur à prédire. Ceci est résumé sur le graphique suivant:

$y(m-p-1, n-q-1)$	$y(m, n-q-1)$
x	x
.	.
.	.
x	x $y(m, n-1)$
$y(m-p-1, n)$	$y(m-1, n)$ $y(m, n)$

Les équations de récurrence (2.3) et (2.4) du modèle 1-D restent valables avec la nouvelle définition de la matrice Y donnée en (2.7). Le modèle 2-D permet plusieurs possibilités de déplacement:

$y(m,n) \rightarrow y(m+1,n)$	déplacement horizontal,
$y(m,n) \rightarrow y(m,n+1)$	déplacement vertical,
$y(m,n) \rightarrow y(m+1,n+1)$	déplacement diagonal.

Le choix dépend, lors de la programmation de l'algorithme, de la seule mise à jour de la matrice Y .

A chaque pixel (m,n) , l'erreur de prédiction $\varepsilon(m,n)$ est calculée: $\varepsilon(m,n) = y(m,n) - z(m,n)$.

L'image composée des erreurs $\varepsilon(m,n)$ constitue le support de l'étude proposée.

En dehors des zones de caractères, les erreurs de prédiction sont connues pour leur gaussiannité. Leur valeur augmente, en valeur absolue, lors d'un changement de modèle.

Du fait de l'hypothèse de gaussiannité des erreurs, un seuil adaptatif peut être choisi, permettant de ne retenir pour l'identification du changement de modèle, que les zones ayant provoqué un dépassement de seuil.

Le seuil α est choisi tel que la probabilité pour que $|\varepsilon(m,n)|$ dépasse le seuil soit par exemple de 1%. Cette valeur correspond à 2.57σ , où σ est l'écart-type, et est récursivement recalculée sur l'image.

3. CONSTRUCTION D'UN HISTOGRAMME

Nous souhaitons approcher la densité de probabilité de la

variable aléatoire $\varepsilon = (\varepsilon_k)$, $1 \leq k \leq N$, des erreurs de prédiction précédentes à partir d'un histogramme. Soit Ω l'ensemble de toutes les valeurs prises par ε et soit λ la vraie loi de probabilité de ε , supposée absolument continue par rapport à une loi a priori notée μ . La densité de probabilité $f(\lambda, \varepsilon)$ de ε s'écrit:

$$\forall \varepsilon \in \Omega \quad f(\lambda, \varepsilon) = \frac{d\lambda}{d\mu}(\varepsilon) \quad (3.1)$$

où $\frac{d\lambda}{d\mu}$ est la dérivée de Radon-Nykodim de λ par rapport à μ et $f(\lambda, \varepsilon)$ la densité de probabilité inconnue. Celle-ci devra être approchée de manière optimale, au sens d'un coût à définir, par un histogramme à c classes construit avec N valeurs ε_k de ε . Trois critères vont être proposés donnant c , pour estimer λ par ce que nous noterons θ_c .

3.1. Critère de sélection du nombre de classes d'un histogramme

Il est connu que l'estimateur θ_M au sens du maximum de vraisemblance de la restriction λ_M de λ à une partition M à m classes est donné par, avec $A_i \in M$:

$$\forall i \in \{1, \dots, m\}, \theta_{M(A_i)} = \frac{\text{nombre d'éléments } \varepsilon_k \text{ dans } A_i}{N} \quad (3.2)$$

Il se confond donc avec la fréquence statistique de A_i et cet estimateur est indépendant de la partition contenant A_i .

Avec cette écriture (3.1) devient:

$$f(\theta_M, \varepsilon) = \prod_{i=1}^m \frac{\theta_{M(A_i)}}{\mu(A_i)} \mathbb{1}_{A_i}(\varepsilon) \quad (3.3)$$

Nous cherchons à déterminer la sous partition C de M à c classes qui minimise une fonction de perte $W(\theta_c, \lambda)$. Pour définir cette fonction, nous utiliserons des fonctions ϕ convexes et positives telles que:

$$W(\theta_c, \lambda) \triangleq E_\lambda \left(\phi \left(\frac{f(\theta_c, \varepsilon)}{f(\lambda, \varepsilon)} \right) \right) \quad (3.4)$$

qui d'après (3.3), et en notant E_λ l'espérance mathématique relativement à la loi λ , s'écrit:

$$W(\theta_c, \lambda) = \sum_{i=1}^c \lambda(B_i) \phi \left(\frac{\theta_c(B_i)}{\lambda(B_i)} \right), \text{ avec } \bigcup_{i=1}^c B_i = C.$$

Par exemple, pour $\phi(x) = -2 \ln x$, la fonction de perte $W(\theta_c, \lambda)$ s'écrit (contraste ou information de Kullback):

$$W(\theta_c, \lambda) = -2 \sum_{i=1}^c \lambda(B_i) \ln \left(\frac{\theta_c(B_i)}{\lambda(B_i)} \right) \quad (3.5)$$

mais dans ce cas, $W(\theta_c, \lambda)$ peut être égal à l'infini avec une probabilité non nulle. Nous lui préférons $\phi(x) = 4(\sqrt{x}-1)^2$ et alors $W(\theta_c, \lambda)$ devient (distance de Hellinger):

$$W(\theta_c, \lambda) = 4 \sum_{i=1}^c \lambda(B_i) \left(\sqrt{\frac{\theta_c(B_i)}{\lambda(B_i)}} - 1 \right)^2 \quad (3.6)$$

Si nous notons λ_c la probabilité relativement à la partition C , si nous notons $|\lambda - \lambda_c|_\infty \triangleq \sup_{A \in M} \left| \left(\frac{\lambda_c(B)\mu(A)}{\mu(B)\lambda(A)} - 1 \right) \right|$ avec B l'unique

élément de C contenant l'élément A de M , si nous notons

$$\sum_{A \in M} \frac{1}{\lambda(A)} [\theta(A) - \lambda(A)]^2 = \|\theta_c - \lambda\|^2, \text{ alors:}$$

Théorème:

Les variables aléatoires $NW(\theta_c, \lambda)$ et $N\|\theta_c - \lambda\|^2$ ont asymptotiquement à α près, la même loi de probabilité sous la conditions $|\lambda - \lambda_c|_\infty < \alpha$. Cette loi est une loi du χ^2 à $(c-1)$ degrés de liberté ■

En réalité, si nous considérons plusieurs échantillons (ε_k) ,

$k \in N$, il paraît naturel de chercher à minimiser le risque $\overline{W} \triangleq E(W(\theta_c, \lambda))$. On peut alors montrer ([ELMA87]) que \overline{W} s'exprime à α près, par $E(\|\theta_c - \lambda\|^2)$. Nous renvoyons le lecteur à [OLIV93] pour la démonstration de cet important théorème d'approximation de la perte W puis de son espérance \overline{W} .

On se retrouve ainsi avec la même approximation du risque \overline{W} que pour l'obtention du critère d'Akaïke classique, mais à partir d'un contexte totalement différent, puisqu'il s'agit d'estimer, dans ce dernier cas, le nombre de paramètres de modèles de type ARMA.

De même, après cette estimation de \overline{W} , et si on utilise le contraste de Kullback, on montre ([ELMA93]), que la technique de Rissanen ([RISS86]), s'applique à notre contexte, tout comme le critère de Hannan et Quinn ([HANN79]), après généralisation. On aboutit ainsi aux trois critères que nous appelons respectivement d'Akaïke, Rissanen et Hannan-Quinn:

$$AIC(c) = -2 \sum_{B \in C} \theta_c(B) \ln \frac{\theta(B)}{\mu(B)} + \frac{2c-1}{N} \quad (3.7)$$

$$AIC^*(c) = -2 \sum_{B \in C} \theta_c(B) \ln \frac{\theta(B)}{\mu(B)} + c \frac{[1 + \ln N]}{N} \quad (3.8)$$

$$\phi^*(c) = -2 \sum_{B \in C} \theta_c(B) \ln \frac{\theta(B)}{\mu(B)} + c \frac{[1 + \ln N \ln N]}{N} \quad (3.9)$$

Les critères AIC^* , ϕ^* tiennent compte des pénalisations d'Akaïke et de Rissanen d'une part, d'Akaïke et de Hannan et Quinn d'autre part.

3.2 Obtention de l'histogramme

On part d'un histogramme à m classes, m de l'ordre de $2\sqrt{N}-1$. La technique utilisée, consiste à regrouper des classes adjacentes de la partition jusqu'à obtention du minimum du critère.

A chaque itération k , on choisit le regroupement parmi les $(k-1)$ possibles, minimisant AIC , AIC^* ou ϕ^* , k allant de m à c , selon le critère adopté ([COLO93] sur AIC). La figure 2 illustre l'évolution de $AIC(k)$ ainsi que le comportement de l'algorithme de fusion de classes pour k allant de $m=30$ à $c=10$.



La loi représentée est celle des erreurs de prédiction obtenue sur 256 pixels de tracé de fond de l'image montrée en figure 1.

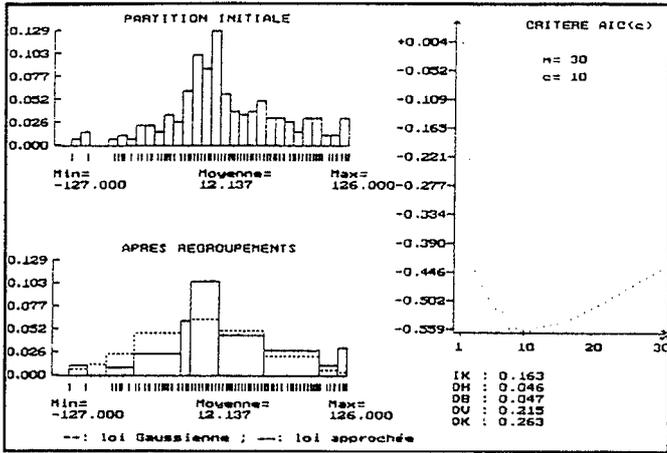


Figure 2: Fusion de classes d'un histogramme sur tracé de fond

4. DETECTION DE CHANGEMENTS DE LOI

Les mesures de dissemblance entre histogrammes testées sont les distances suivantes:

- Hellinger: $H^2(\theta, \nu) = \frac{1}{2} \sum_{i=1}^m (\sqrt{\theta(A_i)} - \sqrt{\nu(A_i)})^2$ (4.1)

- Bhattacharyya: $B(\theta, \nu) = -\text{Ln} \sum_{i=1}^m \sqrt{\theta(A_i)\nu(A_i)}$ (4.2)

- Kolmogorov: $V(\theta, \nu) = \sum_{i=1}^m |\theta(A_i) - \nu(A_i)|$ (4.3)

Après dépassement du seuil α (cf. 2.2), l'histogramme des erreurs de prédiction est construit sur une fenêtre de N pixels, et comparée à l'histogramme de référence du tracé de fond réactualisé avant rupture. Ce dernier histogramme est préféré à celui construit à partir d'une loi gaussienne a priori, mais peut entraîner des classes vides (voir figure 2), ce qui exclut l'utilisation des information et divergence de Kullback, celles-ci pouvant être alors infinies.

5. RESULTATS ET DISCUSSION

L'évolution des critères proposés est montrée en figure 3, sur une zone de caractères de l'image de la figure 1.

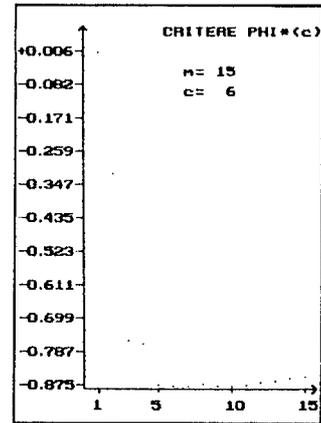
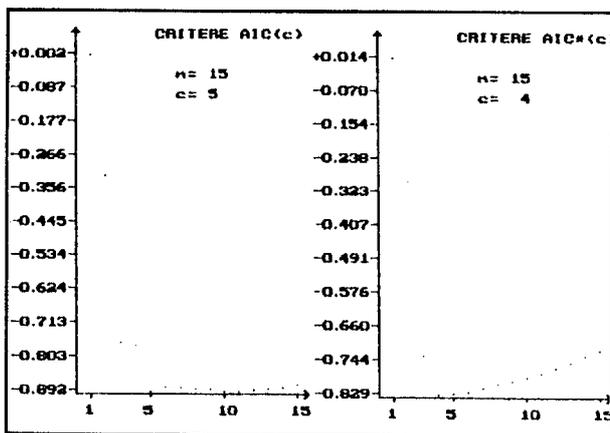


Figure 3: Evolution des critères AIC(c), AIC*(c) et $\phi^*(c)$

Sur l'ensemble des zones testées sur différentes images, il s'avère que le critère AIC* présente le minimum le plus marqué, bien que les trois valeurs de c obtenues soient égales à une unité près, en accord avec les expressions (3.7) à (3.9).

Aussi, ce sont les histogrammes obtenus par AIC* qui sont retenus dans notre étude, comme estimation des lois d'innovation correspondant aux zones éventuelles de caractères. Ils sont comparés aux histogrammes de référence (cf. §4) par les trois mesures H², V et B. Sur les zones de caractères des images, ces trois mesures présentent des valeurs numériques plus élevées permettant d'établir un seuil de segmentation à l'heure actuelle empirique.

REFERENCES

[COLO93]: Colot O., "Apprentissage et détection automatique de changements de modèles - Application aux signaux électroencéphalographiques", Thèse Rouen, 1993.

[COUR92]: Courtellemont P., Olivier C. & Avila M., "2-D Modelling and Hough Transform for Document Analysis.", *Proc. of The IEEE-IES, Iecon'92*, Vol. 3, San Diego, California, USA, pp. 1325-1330.

[ELMA87]: El Matouat A.: "Sélection du nombre de paramètres d'un modèle. Comparaison avec le critère d'Akaïke.", Thèse Rouen, 1987.

[ELMA93]: El Matouat A., "Ordre d'un modèle et contraste de Kullback.", Thèse ULB, 1993.

[HANN79]: Hannan E.J. & Quinn B.G., "The Determination of the Order of an Autoregression.", *Jal Royal Stat. Soc.*, B, 41, 2, 1979, pp. 190-195.

[OLIV93]: Olivier C, Courtellemont P., Colot O., de Bruçq D. & El Matouat A., "Comparison of Histogramms: A Tool For Detection.", *Proc. of The Int. Conf. on Fault Diagnosis*, Toulouse 5-7 Avril 1993, Vol. 2, pp. 610-614.

[RISS86]: Rissanen J., "Stochastic Complexity and the Maximum Entropy Principle.", *The Annals of Statistics*, Vol. 14, 3, 1986, pp. 1080-1100.